



## RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/668,749  
Source: DIPE  
Date Processed by STIC: 10/7/2003

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 4.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker/chkr41note.htm>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry directly to (EFFECTIVE 12/01/2003):  
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 10/08/2003

## Raw Sequence Listing Error Summary

ERROR DETECTED	SUGGESTED CORRECTION	SERIAL NUMBER: <u>10/668,749</u>
ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE		
1 _____ Wrapped Nucleics Wrapped Aminos	The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."	
2 _____ Invalid Line Length	The rules require that a line <b>not exceed</b> 72 characters in length. This includes white spaces.	
3 _____ Misaligned Amino Numbering	The numbering under each 5 <sup>th</sup> amino acid is misaligned. Do <b>not</b> use tab codes between numbers; use <b>space characters</b> , instead.	
4 _____ Non-ASCII	The submitted file was <b>not</b> saved in ASCII(DOS) text, as <b>required</b> by the Sequence Rules. Please <b>ensure your subsequent submission is saved in ASCII text</b> .	
5 _____ Variable Length	Sequence(s) _____ contain n's or Xaa's representing more than one residue. <b>Per Sequence Rules, each n or Xaa can only represent a single residue.</b> Please present the <b>maximum</b> number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.	
6 _____ PatentIn 2.0 "bug"	A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. <b>This applies to the mandatory &lt;220&gt;-&lt;223&gt; sections for Artificial or Unknown sequences.</b>	
7 _____ Skipped Sequences (OLD RULES)	Sequence(s) _____ missing. If intentional, please insert the following lines for <b>each</b> skipped sequence: (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading) (xi) SEQUENCE DESCRIPTION: SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) This sequence is intentionally skipped  Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to <b>include</b> the skipped sequences.	
8 _____ Skipped Sequences (NEW RULES)	Sequence(s) _____ missing. If <b>intentional</b> , please insert the following lines for <b>each</b> skipped sequence. <210> sequence id number <400> sequence id number 000	
9 _____ Use of n's or Xaa's (NEW RULES)	Use of n's and/or Xaa's have been detected in the Sequence Listing. Per 1.823 of Sequence Rules, use of <220>-<223> is <b>MANDATORY</b> if n's or Xaa's are present. In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.	
10 _____ Invalid <213> Response	Per 1.823 of Sequence Rules, the only <b>valid</b> <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is <b>required</b> when <213> response is Unknown or is Artificial Sequence	
11 _____ Use of <220>	Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is <b>MANDATORY</b> if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section. (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)	
12 _____ PatentIn 2.0 "bug"	Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use <u>File Manager</u> or any other manual means to copy file to floppy disk.	
13 _____ Misuse of n/Xaa	"n" can only represent a single <u>nucleotide</u> ; "Xaa" can only represent a single <u>amino acid</u>	



IFWO

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/668,749

DATE: 10/07/2003  
TIME: 09:18:34

Input Set : A:\50112-1580.txt  
Output Set : N:\CRF4\10072003\J668749.raw

3 <110> APPLICANT: Agilent Technologies  
5 <120> TITLE OF INVENTION: Methods and Systems for Nanopore Data Analysis  
7 <130> FILE REFERENCE: 50112-1580  
C--> 9 <140> CURRENT APPLICATION NUMBER: US/10/668,749

10 <141> CURRENT FILING DATE: 2003-09-23  
12 <160> NUMBER OF SEQ ID NOS: 8  
14 <170> SOFTWARE: PatentIn version 3.2

16 <210> SEQ ID NO: 1  
17 <211> LENGTH: 1300  
18 <212> TYPE: DNA  
19 <213> ORGANISM: neucleotides  
21 <400> SEQUENCE: 1

*Invalid response - see  
item 10 on  
Erra summary sheet*

Does Not Comply  
Corrected Diskette Needed

22	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	60
24	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	120
26	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	180
28	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	240
30	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	300
32	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	360
34	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	420
36	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	480
38	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	540
40	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	600
42	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	660
44	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	720
46	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	780
48	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	840
50	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	900
52	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	960
54	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1020
56	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1080
58	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1140
60	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1200
62	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1260
64	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1300

67 <210> SEQ ID NO: 2  
68 <211> LENGTH: 500  
69 <212> TYPE: DNA  
70 <213> ORGANISM: neucleotides  
72 <400> SEQUENCE: 2

73	cccccccccc	cccccccccc	cccccccccc	cccccccccc	cccccccccc	cccccccccc	60
75	cccccccccc	cccccccccc	cccccccccc	cccccccccc	cccccccccc	cccccccccc	120
77	cccccccccc	cccccccccc	cccccccccc	cccccccccc	cccccccccc	cccccccccc	180
79	cccccccccc	cccccccccc	cccccccccc	cccccccccc	cccccccccc	cccccccccc	240

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/668,749

DATE: 10/07/2003

TIME: 09:18:34

Input Set : A:\50112-1580.txt

Output Set: N:\CRF4\10072003\J668749.raw

```
81 cccccccccc cccccccccc cccccccccc cccccccccc cccccccccc cccccccccc 300
83 cccccccccc cccccccccc cccccccccc cccccccccc cccccccccc cccccccccc 360
85 cccccccccc cccccccccc cccccccccc cccccccccc cccccccccc cccccccccc 420
87 cccccccccc cccccccccc cccccccccc cccccccccc cccccccccc cccccccccc 480
89 cccccccccc cccccccccc 500
92 <210> SEQ ID NO: 3
93 <211> LENGTH: 70
94 <212> TYPE: DNA
95 <213> ORGANISM: neucleotides
97 <400> SEQUENCE: 3
98 ccacaaacaa acaaccacac aaacacacaa ccacaacacc aacacacaaa caaaccaaca 60
100 cacaaactcc 70
103 <210> SEQ ID NO: 4
104 <211> LENGTH: 48
105 <212> TYPE: DNA
106 <213> ORGANISM: neucleotides
108 <400> SEQUENCE: 4
109 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 48
112 <210> SEQ ID NO: 5
113 <211> LENGTH: 196
114 <212> TYPE: DNA
115 <213> ORGANISM: neucleotides
117 <400> SEQUENCE: 5
118 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 60
120 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 120
122 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 180
124 aaaaaaaaaa aaaaaa 196
127 <210> SEQ ID NO: 6
128 <211> LENGTH: 48
129 <212> TYPE: DNA
130 <213> ORGANISM: neucleotides
132 <400> SEQUENCE: 6
133 caaacaaacc aacacacaaa ctcccctcaa acacacaacc aaacaaac 48
136 <210> SEQ ID NO: 7
137 <211> LENGTH: 100
138 <212> TYPE: DNA
139 <213> ORGANISM: neucleotides
141 <400> SEQUENCE: 7
142 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 60
144 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 100
147 <210> SEQ ID NO: 8
148 <211> LENGTH: 87
149 <212> TYPE: DNA
150 <213> ORGANISM: neucleotides
152 <400> SEQUENCE: 8
153 ccacaaacaa acaaccacac aaacacacaa ccacaacacc aacacacaaa caaaccaaca 60
155 cacaaactcc tatagtgagt cgtatta 87
```

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/668,749

DATE: 10/07/2003

TIME: 09:18:35

Input Set : A:\50112-1580.txt

Output Set: N:\CRF4\10072003\J668749.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application Number